CLAIMS

WHAT IS CLAIMED IS:

1. A humming transcription system comprising:

an humming signal input interface accepting an input humming signal; and

a humming transcription block that transcribes the input humming signal into a musical sequence, wherein the humming transcription block includes a note segmentation stage that segments note symbols in the input humming signal based on note models defined by a note model generator, and a pitch tracking stage that determines the pitches of the note symbols in the input humming signal based on pitch models defined by a statistical model.

2. The humming transcription system of claim 1 further comprising a humming database recording a sequence of humming data provided to train the note models and the pitch models.

3. The humming transcription system of claim 1 wherein the note model generator is implemented by phone-level Hidden Markov Models with Gaussian Mixture Models.

4. The humming transcription system of claim 3 wherein the phone-level Hidden Markov Models further comprising a

silence model for preventing errors of segmenting the note symbols in the input humming signal caused by noises and signal distortions imposed on the input humming signal.

5. The humming transcription system of claim 3 wherein the phone-level Hidden Markov Models define the note models based on a feature vector associated with the characterization of the note symbols in the humming signal, and wherein the feature vector is extracted from the humming signal.

6. The humming transcription system of claim 5 wherein the feature vector is constituted by at least one Mel-Frequency Cepstral Coefficient, an energy measure, and first-order derivatives and second-order derivatives thereof.

7. The humming transcription system of claim 1 wherein the note segmentation stage further includes:

a note decoder that recognizes each note symbol in the humming signal; and

a duration model that detects the duration associated with each note symbol in the humming signal and labels the duration of each note symbol relative to a previous note symbol.

8. The humming transcription system of claim 7 wherein the note decoder utilizes a Viterbi decoding algorithm to recognize each note symbol.

9. The humming transcription system of claim 1 wherein the note model generator utilizes a maximum likelihood method with Baum-Welch re-estimation formula to train the note models.

10. The humming transcription system of claim 1 wherein the statistical model is implemented by a Gaussian Model.

11. The humming transcription system of claim 1 wherein the pitch tracking stage further comprising a pitch detector that analyzes the pitch information of the input humming signal, extracts features used to characterize a melody contour of the input humming signal, and detects the relative pitch of the note symbols in the humming signal based on the pitch models.

12. The humming transcription system of claim 11 wherein the pitch detector uses a short-time autocorrelation algorithm to analysis the pitch information of the input humming signal.

13. The humming transcription system of claim 1 further comprising a music language model that predict the current note symbol based on previous note symbols in the musical sequence.

14. The humming transcription system of claim 13 wherein the music language model is implemented by a N-gram duration model that predicts the relative duration associated with the current note symbol based on relative durations associated with previous note symbols in the musical sequence.

15. The humming transcription system of claim 13 wherein the music language model includes a N-gram pitch model that predicts the relative pitch associated with the current note symbol based on relative pitches associated with previous note symbols in the musical sequence.

16. The humming transcription system of claim 13 wherein the music language model includes a N-gram pitch and duration model that predicts the relative duration associated with the current note symbol based on relative durations associated with previous note symbols in the musical sequence, and predicts the relative pitch

associated with the current note symbol based on relative pitches associated with previous note symbols in the musical sequence.

17. The humming transcription system of claim 1 wherein the humming transcription system is arranged in a computing machine.

18. A humming transcription methodology comprising:

compiling a humming database recording a sequence of humming data;

inputting a humming signal;

segmenting the humming signal into note symbols according to note models defined by a note model generator; and

determining the pitch value of the note symbols based on pitch models defined by a statistical model.

19. The humming transcription methodology of claim 18 wherein segmenting the humming signal into note symbols includes the steps of:

extracting a feature vector comprising a plurality of features used to characterize the note symbols in the humming signal;

defining the note models based on the features vector;

recognizing each note symbol in the humming signal based on an audio decoding method by using the note models; and

labeling the relative duration of each note symbol in the humming signal.

20. The humming transcription methodology of claim 19 wherein the note model generator is implemented by phone-level Hidden Markov Models incorporating a silence model with Gaussian Mixture Models.

21. The humming transcription methodology of claim 19 wherein the feature vector is extracted from the humming signal.

22. The humming transcription methodology of claim 19 wherein the note models are trained by using the humming data extracted from the humming database.

23. The humming transcription methodology of claim 19 wherein the audio decoding method is a Viterbi decoding algorithm.

24. The humming transcription methodology of claim 18 wherein determining the pitch value of each note symbol includes the steps of:

analyzing the pitch information of the input humming signal;

extracting features used to build a melody contour of the humming signal; and

detecting the relative pitch interval of each note symbol in the input humming signal based on the pitch models.

25. The humming transcription methodology of claim 24 wherein analyzing the pitch information of the input humming signal is accomplished by using a short-time autocorrelation algorithm.

26. The humming transcription methodology of claim 18 wherein the statistical model is a Gaussian model.